

**IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
SHERMAN DIVISION**

THE STATE OF TEXAS, et al.,

Plaintiffs,

v.

GOOGLE LLC,

Defendant.

§
§
§
§
§
§
§

Civil Action No. 4:20-cv-00957-SDJ

**GOOGLE LLC'S REPLY IN SUPPORT OF ITS MOTION TO EXCLUDE
EXPERT TESTIMONY OF JACOB HOCHSTETLER**

Table of Contents

I. HOCHSTETLER’S LOG DATASET OPINIONS ARE IRRELEVANT, UNHELPFUL, AND UNFAIRLY PREJUDICIAL.....	1
II. HOCHSTETLER IS NOT QUALIFIED TO GIVE THE GENERALIZED OPINIONS; THEY ARE UNRELIABLE, IRRELEVANT AND PREJUDICIAL.....	2

TABLE OF AUTHORITIES

Cases	Page(s)
<i>Aguilar v. Immigration & Customs Enft Div. of U.S. Dep't of Homeland Sec.</i> , 255 F.R.D. 350 (S.D.N.Y. 2008)	4
<i>Bunch v. Bullard</i> , 795 F.2d 384 (5th Cir. 1986)	5
<i>Charles v. K-Patents, Inc.</i> , 2020 WL 1480734 (E.D. Tex. Mar. 19, 2020)	5
<i>Hathaway v. Bazany</i> , 507 F.3d 312 (5th Cir. 2007)	3
<i>Jahn v. Equine Services, PSC</i> , 233 F.3d 382 (6th Cir. 2000)	4
<i>Lyondell Chem. Co. v. Albemarle Corp.</i> , 2007 WL 9711338 (E.D. Tex. May 31, 2007).....	5
<i>McDaniels v. Ross</i> , 2020 WL 5901120 (N.D. Tex. Feb. 6, 2020).....	2
<i>Reid v. Albemarle Corp.</i> , 207 F. Supp. 2d 499 (M.D. La. 2001).....	3
<i>Rimkus Consulting Group, Inc. v. Cammarata</i> , 688 F. Supp. 2d 598 (S.D. Tex. 2010)	1
<i>United States v. Rider</i> , 2022 WL 2813719 (E.D. Tex. July 18, 2022)	2
<i>Van Winkle v. Rogers</i> , 82 F.4th 370 (5th Cir. 2023)	3
Other Authorities	
Fed. R. Evid. 702	1, 2, 5

Rather than squarely addressing the irrelevance or unreliability of Hochstetler’s opinions, Plaintiffs’ Response to Google’s Motion to Exclude Jacob Hochstetler’s Chat-related opinions (“Response”) tries to shift blame for their shortcomings onto Google, for providing only a “tiny sliver of log metadata” about Chat messages. Resp. (ECF 784) at 1. Putting aside the irrelevance of *metadata about* Chat usage—which reveals no message content but simply when history was “on” or “off”—it was **Plaintiffs** who wrote their own fate. Despite learning **five years ago** about Google’s Chat retention process, they stayed quiet for years, never inquiring about the existence of Chat-related metadata. They waited until **2024** to seek (only) the metadata extract Google had produced a year earlier in *Play*¹ (which case had nothing to do with display advertising), created from logs that—by their nature—only contain two months of data. Plaintiffs’ own inaction, coupled with the irrelevance and unreliability of Hochstetler’s Chat-related opinions, set the stage for exclusion under Rules 702 and 403.

I. HOCHSTETLER’S LOG DATASET OPINIONS ARE IRRELEVANT, UNHELPFUL, AND UNFAIRLY PREJUDICIAL.

The Log Dataset Opinions declare that five high-level Google employees whose day-to-day jobs were *not* in display advertising sent ~20,000 messages per year with “history off.” But Rule 702(a) requires that expert testimony illuminate a “fact in issue.” Mot. at 3–4; Resp. at 6. The only fact “in issue” is whether Google took reasonable steps to preserve evidence “**relevant** to the party’s **claim or defense**.”² And Plaintiffs agree that Hochstetler offers no opinion on whether unsaved messages were relevant to this case. Resp. at 7. To obscure this irrelevance, Plaintiffs raise a smokescreen of “facts *not* in issue.” They exalt the technical nature of Hochstetler’s

¹ All capitalized terms have the same meaning as in Google’s Motion (ECF 762). Unless otherwise indicated, quotations and citations have been omitted for ease of reading, and all emphasis is added.

² *Rimkus Consulting Group, Inc. v. Cammarata*, 688 F. Supp. 2d 598, 615–16 (S.D. Tex. 2010); *see also* ECF 757 at 8–20.

analysis, *id.* at 6–7, but Google is not challenging the *reliability* of his message counts *within* the Log Dataset, Mot. at 6-9. The issue is a lack of *relevance*, substantially outweighed by the unfair prejudice from allowing him to testify about Chat.³

As for that prejudice, Plaintiffs do not address the problem of jurors being swayed by the “mystique” of expert testimony and struggling to parse Hochstetler’s two separate technical opinions and relevant qualifications. *See* Mot. at 8; *United States v. Rider*, 2022 WL 2813719, at *2 (E.D. Tex. July 18, 2022), *aff’d*, 94 F.4th 445 (5th Cir. 2024) (excluding irrelevant expert opinion; “because of an expert’s stature *qua* expert, jurors may assign more weight to expert testimony than it deserves”).⁴ And Plaintiffs completely ignore the prejudice from introducing the *Play* case, as well as Hochstetler’s misleading emphasis on Google’s CEO.

II. HOCHSTETLER IS NOT QUALIFIED TO GIVE THE GENERALIZED OPINIONS; THEY ARE UNRELIABLE, IRRELEVANT AND PREJUDICIAL.

Hochstetler’s Generalized Opinions—which Plaintiffs call “Deletion Opinions”—are confined to a single paragraph and footnote in his report. ECF 762-2 ¶ 82 & n.106 (using multiplication to conclude that if 5 employees sent 18,566 “history off” messages over 68 days, then 141 employees would have sent “close to 1.5 million” annually). Plaintiffs distract from this rudimentary math by jumbling Hochstetler’s Log Dataset analysis with his Generalized Opinions, but these “extrapolations” are ones that he is unqualified to make and that lack any statistical basis

³ Plaintiffs also claim the Log Dataset Opinions are helpful because Plaintiffs’ other Chat-related expert, Grande, relied on them. Resp. at 7–8. But this only makes them relevant *to Grande*, not to any fact “in issue”—Rule 702’s standard. Besides, Grande’s opinions rest on Hochstetler’s *Generalized Opinions* (regarding 141 to 202 Google employees over several years) not his narrower *Log Dataset Opinions* (regarding 5 Google employees over 68 days). *See* ECF 764-5.

⁴ Plaintiffs mischaracterize *McDaniels v. Ross* on this point. 2020 WL 5901120 (N.D. Tex. Feb. 6, 2020). The movants in that case *never challenged the expert* based on confusion from discrete topics of testimony. *See* Mot. to Exclude Pl. McDaniels’ Expert, Jeff G. Vick [ECF 131] at 13 (No. 2:17-00240) (Aug. 30, 2019). Instead, their prejudice argument centered on whether “negligence per se” had been pleaded. *Id.*

or methodology.

Hochstetler is Not Qualified. He has no statistics background, disclaimed basic statistical knowledge, and confirmed he has never “represented [him]self as an expert in statistical analysis”—including for this engagement. Mot. at 9-11.⁵ To wriggle out of this, Plaintiffs pursue two (contradictory) arguments. First, they say statistics is irrelevant because Hochstetler analyzed the full Log Dataset. Resp. at 10. But this conflates his Log Dataset Opinions with his Generalized Opinions, which extrapolate far beyond the Log Dataset. Extrapolating from a smaller sample to draw broader conclusions *is* statistics.⁶ Second, they claim he *is* qualified in statistics, pointing to a single professional paper (Resp. at 11)—but that involved writing code to understand a large, complete dataset (computer science), ***not*** calculating broad conclusions from an incomplete sample (statistics).⁷ Although experience can qualify an expert, Hochstetler has not “produced sufficient information about that work to demonstrate his knowledge of statistical science in this case.” *Reid v. Albemarle Corp.*, 207 F. Supp. 2d 499, 501 (M.D. La. 2001).

Hochstetler’s Opinions Are Not Reliable. “[R]eliable methodology is in all instances mandatory.” *Hathaway v. Bazany*, 507 F.3d 312, 318 (5th Cir. 2007). The standard is not—as Plaintiffs would have it—“as reliable as possible” or the “best possible estimate.” Resp. at 10. Hochstetler did nothing resembling statistical analysis and did not even examine whether his sample of five employees over 68 days in late 2022–23 was representative—which it was not. Mot.

⁵ In the very case Plaintiffs rely on, Resp. at 4, the Fifth Circuit upheld the *exclusion* of certain expert opinions when the expert “himself admitted his lack of qualification and expertise” in that specific area. *Van Winkle v. Rogers*, 82 F.4th 370, 380 (5th Cir. 2023). Just so here.

⁶ Plaintiffs also claim that Hochstetler’s Generalized Opinions are not “statistics” because they are an “opinion on extrapolation from large datasets.” Resp. at 11. But Hochstetler did not take a large dataset and extrapolate a smaller conclusion—he did ***just the opposite***.

⁷ ECF 762-4 at 23:10-24:10 (describing the paper), 25:14–19 (Q. [B]esides this smart city planning paper that LA County did not implement, have you ever at any other point been hired to perform a statistical analysis?...A. No.).

at 12–13.⁸ That Hochstetler saw no evidence “to the contrary,” Resp. at 13, means only that he turned a blind eye to evidence—not that his sample *was* representative.⁹

Admitting the “paucity in the data,” Resp. at 1, Plaintiffs blame Google for not providing more Chat *metadata*. This does not push Hochstetler’s Generalized Opinions over the reliability line. First, Google did not “destroy[]” any Chat metadata. Resp. at 1, 4 n.1. As recognized by Hochstetler (but Plaintiffs’ Response ignores), Chat logs are written over every few months because they are kept only for *debugging* purposes. ECF 762 at 2.¹⁰ Second, any “paucity” of additional, potentially **relevant** log metadata falls at Plaintiffs’ feet. As early as 2020, when Google first disclosed its preservation practices,¹¹ Plaintiffs could have inquired about the existence of Chat logs/metadata, and/or requested then-existing logs for certain ad-tech employees. ***They did neither.*** They instead waited until 2024 to seek (only) the back-end logs that Google had produced in *Play*. ECF 348-2 (RFP No. 133, served 4/3/24).¹²

Third, Plaintiffs claim—without authority—that an unreliable expert opinion automatically becomes admissible if the opposing party (at some point) had more data that might have helped the analysis satisfy Rule 702. But in *Charles v. K-Patents, Inc.*, for instance (Resp. at 13), the court

⁸ Plaintiffs acknowledge the importance of representativeness: their Motion to Exclude Google’s survey expert Simonson argues that his sample was unrepresentative. ECF 673 at 9–13.

⁹ See ECF 762-4 at 99:10–15 (had “no info about the 136 [employees] outside of their names.”).

¹⁰ To be clear, Plaintiffs do not allege that Google spoliated (in the face of a duty to preserve) Chat ***metadata*** (the back-end logs), but rather, Chat messages themselves. See ECF 693 (Mot. for Spoliation Sanctions). Plaintiffs’ reliance on *Jahn v. Equine Services, PSC*, 233 F.3d 382 (6th Cir. 2000)—a case where an operation allegedly caused a horse’s death, but the defendant veterinary hospital failed to keep any post-operative medical records—is misplaced.

¹¹ See ECF 757 (Google Resp. to Mot. for Spoliation Sanctions) at 3–4.

¹² Cf. *Aguilar v. Immigration & Customs Enf’t Div. of U.S. Dep’t of Homeland Sec.*, 255 F.R.D. 350, 364 (S.D.N.Y. 2008) (“This lawsuit demonstrates why it is so important that parties fully discuss their ESI early in the evolution of a case. Had that been done, the Defendants might not have opposed the Plaintiffs’ requests for certain metadata.... Instead, these proceedings have now been bogged down in expensive and time-consuming litigation of electronic discovery issues only tangentially related to the underlying merits....”).

merely held that when sufficient data is lacking to form a “precise” conclusion, he “may opine as to a reasonable estimate, *as long as the underlying methodology is reliable*.” 2020 WL 1480734, at *5 (E.D. Tex. Mar. 19, 2020), *aff’d*, 2021 WL 826402 (5th Cir. Mar. 3, 2021) (emphasis added). And the court—in *excluding* much of the expert’s testimony—cautioned that “[a]n expert may not, however, simply hazard an unsubstantiated guess masked as technical expertise” and must “offer good reason to think that his approach produces an accurate estimate using professional methods.” *Id.* at *5. Here, Plaintiffs took numerous depositions, received millions of emails and individual Chat messages, and propounded dozens of interrogatories. Hochstetler analyzed *none* of this to test the representativeness of the Log Dataset before “extrapolating” beyond it.¹³

Hochstetler’s Opinions are irrelevant, unhelpful, and unduly prejudicial. Like his Log Dataset Opinions, Hochstetler’s Generalized Opinions do not address whether a single one of the “millions” of “history off” messages was relevant. And Plaintiffs showcase the exact undue prejudice justifying exclusion: their Response muddles Hochstetler’s opinions, implying his Generalized Opinions are “rooted in computer science,” span “42 pages” of his report and involve “custom source code.” Resp. at 15. That describes his *Log Dataset Opinions*, not his broader Generalized Opinions—derived by simple math shown in a footnote. ECF 762-2 ¶ 82 & n.106. This is precisely the type of confusion a jury would experience as the lines blur between Hochstetler’s opinions and expertise (or lack thereof) in two separate areas.

¹³ In *Bunch v. Bullard*, 795 F.2d 384 (5th Cir. 1986) (cited in Resp. at 10), the issue was not the admissibility of expert testimony under *Daubert* or Rule 702, but whether an employer could evade liability for race discrimination by citing its workforce’s small size. *Id.* at 395. In *Lyondell Chem. Co. v. Albemarle Corp.*, 2007 WL 9711338 (E.D. Tex. May 31, 2007) (Resp. at 14), the court’s reluctance to exclude “***otherwise sound methodology*** because of inherent constraints in the underlying evidentiary record” involved decades-old missing records of a defendant’s toxic waste dumping from the 1970s. *Id.* at *4; 2006 WL 8441057, at *21 (E.D. Tex. Apr. 18, 2006) (relevant factual findings). No such “sound methodology” or “inherent constraints” exist here. Notably, Plaintiffs cite cases involving incomplete records ***on the merits***—not discovery-on-discovery.

Dated: February 7, 2025

Respectfully submitted,

/s/ Kathy D. Patrick

Kathy D. Patrick
State Bar No. 15581400
KPatrick@gibbsbruns.com
Ayesha Najam
State Bar No. 24046507
ANajam@gibbsbruns.com
Michael Doman
State Bar No. 24131182
MDoman@gibbsbruns.com
Elisa Wulfsberg
State Bar No. 24125469
EWulfsberg@gibbsbruns.com
Gibbs & Bruns LLP
1100 Louisiana, Suite 5300
Houston, Texas 77002
Tel.: (713) 650-8805
Fax: (713) 750-0903

Eric Mahr (*pro hac vice*)
FRESHFIELDS US LLP
700 13th Street, NW
10th Floor
Washington, DC 20005
Telephone: (202) 777-4545
Email: eric.mahr@freshfields.com

Robert J. McCallum (*pro hac vice*)
FRESHFIELDS US LLP
3 World Trade Center
175 Greenwich Street
New York, NY 10007
Telephone: (212) 284-4910
Email: rob.mccallum@freshfields.com

Attorneys for Defendant Google LLC

CERTIFICATE OF SERVICE

I certify that on February 7, 2025, this document was filed electronically in compliance with Local Rule CV-5(a) and served on all counsel who have consented to electronic service, per Local Rule CV-5(a)(3)(A).

/s/ Kathy D. Patrick
Kathy D. Patrick